

Identification of the regulatory genes required for the acid
activation of the low pH inducible gene *lpiA* in *Sinorhizobium*
medicae

Rui, Tian

B.Sc. Molecular Biology (Hons)

This thesis is presented for the degree of

Doctor of Philosophy

Centre for *Rhizobium* studies

Murdoch University

2012

Declaration

I hereby declare that, unless otherwise stated, the work presented in this
thesis is my own.

Rui, Tian

献给我的父母，为你们对我的慈爱和包容。

慈母手中线，
游子身上衣。
临行密密缝，
意恐迟迟归。
谁言寸草心，
报得三春晖。

孟郊. 游子吟

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	I
LIST OF PUBLICATIONS	II
ABBREVIATIONS.....	III
ABSTRACT.....	V
CHAPTER 1 INTRODUCTION	1
1.1 Nitrogen Fixation	2
1.2 Biological Nitrogen Fixation	2
1.3 Symbiotic Nitrogen Fixation	3
1.4 The root nodule bacteria	4
1.5 The rhizobia-legume symbiosis.....	5
1.5.1 The molecular dialogue between rhizobia and legumes.....	6
1.5.2 Host specificity within the rhizobia-legume symbioses	8
1.6 Classification of rhizobia	9
1.7 The genus <i>Sinorhizobium</i>	10
1.7.1 <i>S. meliloti</i>	11
1.7.2 <i>S. medicae</i>	12
1.8 Agriculturally important <i>Medicago</i> hosts for <i>Sinorhizobium</i>	14

1.8.1	<i>Medicago murex</i> L	14
1.8.2	<i>Medicago polymorpha</i> L	15
1.8.3	<i>Medicago sativa</i> L.....	15
1.8.4	<i>Medicago truncatula</i> Gaertn	15
1.9	<i>Sinorhizobium</i> genes required for nodulation and nitrogen fixation	16
1.10	Regulation of symbiotic nitrogen fixation genes in <i>Sinorhizobium</i>	19
1.11	<i>Sinorhizobium</i> sigma factors that regulate symbiotic gene expression	22
1.11.1	RpoN.....	22
1.11.2	RpoH.....	27
1.12	The impact of soil acidity on <i>Sinorhizobium</i> and the <i>Sinorhizobium-Medicago</i> symbiosis ...	28
1.13	The acid-tolerant strain <i>S. medicae</i> WSM419	29
1.14	Identification of genes involved in acid tolerance and response in <i>S. medicae</i> WSM419.....	31
1.15	The genome structure of <i>S. medicae</i> WSM419	34
1.16	Characterisation of the <i>lpiA</i> and its gene region on pSMED02.....	35
1.16.1	Putative function of <i>LpiA</i>	40
1.16.2	FsrR is a regulator of <i>lpiA</i> expression	42
1.17	Research aims.....	43

CHAPTER 2 INACTIVATION OF ACVB, TCSA, AND LPIA IN *S. MEDICAE*

	WSM419.....	45
2.1	Introduction.....	46

2.2 Materials and Methods.....	50
2.2.1 Bacterial strains and plasmids	50
2.2.2 Primer design and PCR amplification	50
2.2.3 PCR product purification	53
2.2.4 Agarose gel electrophoresis	53
2.2.5 Ligation and transformation	54
2.2.6 Plasmid DNA purification by alkaline lysis	54
2.2.7 Biparental conjugation	55
2.2.8 Genomic DNA extraction	56
2.2.9 PCR verification of single crossover mutations in <i>S. medicae</i>	57
2.2.10 DNA sequencing.....	57
2.3 Results.....	59
2.3.1 Methodology of single crossover mutagenesis	59
2.3.2 PCR amplification of <i>tcsA</i> , <i>acvB</i> and <i>lpiA</i> intragenic fragments.....	61
2.3.3 Screening of pGEM-T recombinants	61
2.3.4 Cloning the cassette CAS-1116 into pTR _{tcsA} , pTR _{acvB} and pTR _{lpiA}	66
2.3.5 Removal of pGEM-T sequence from pTR1116 _{tcsA} , pTR1116 _{acvB} and pTR1116 _{lpiA}	68
2.3.6 Mobilization of plasmid to inactivate the target gene into <i>S. medicae</i> WSM419	71
2.3.7 PCR verification of <i>tcsA</i> , <i>acvB</i> , <i>lpiA</i> , <i>tcrA</i> and <i>rpoN</i> single crossover mutations	71
2.3.8 Sequencing PCR verification products of <i>tcsA</i> , <i>tcrA</i> , <i>acvB</i> , <i>lpiA</i> and <i>rpoN</i> mutants.....	73
2.4 Discussion	75

CHAPTER 3 PHENOTYPIC ANALYSIS OF *S. MEDICAE* WSM419

MUTANTS.....77

3.1 Introduction.....	78
------------------------------	-----------

3.2 Materials and Methods.....	81
3.2.1 Bacterial strains and plasmids	81
3.2.2 Culture preparation for phenotype test	82
3.2.3 Stress tolerance tests	82
3.2.4 Antibiotics tests.....	82
3.2.5 Mean generation time (MGT) determination	83
3.2.6 Glasshouse trial	83
3.2.6.1 Experimental design.....	83
3.2.6.2 Preparation of <i>Medicago</i> inoculants.....	84
3.2.6.3 Pot preparation	84
3.2.6.4 Seedlings	84
3.2.6.5 Inoculation of plants and watering schedule	85
3.2.6.6 Harvesting the plants.....	85
3.2.7 Statistical analysis	85
3.3 Results	87
3.3.1 Stress tolerance tests	87
3.3.2 Antibiotic tests	89
3.3.3 Mean generation time (MGT) determination	91
3.3.4 Growth of <i>Medicago</i> in glass house experiments	91
3.3.5 Effect of inoculation on nodulation of hosts	92
3.3.6 Effect of inoculation on dry weight of hosts	94
3.4 Discussion	98

CHAPTER 4 GENE EXPRESSION STUDIES.....103

4.1 Introduction.....	104
------------------------------	------------

4.2 Materials and Methods	108
4.2.1 Bacterial strains and plasmids	108
4.2.2 Triparental conjugation for Cre-mediated excision	108
4.2.3 Selection of CAS-1116 excision mutants	110
4.2.4 PCR verification of Cre-mediated <i>loxP</i> at sites excision	110
4.2.5 Mobilization of pWR220-101H into CAS-1116 excised mutants	110
4.2.6 Measurement of GUS activity	111
4.2.7 RNA extraction	113
4.2.8 Spectrophotometric quantification of RNA	114
4.2.9 RNA formaldehyde-agarose gel electrophoresis	114
4.2.10 Primers design for qRT-PCR	115
4.2.11 First strand cDNA synthesis	115
4.2.12 Quantitative real-time polymerase chain reaction (qRT-PCR)	116
4.2.13 Mathematical model for qRT-PCR data analysis	116
4.3 Results	117
4.3.1 Excision of CAS-1116 from <i>S. medicae</i> mutants using the Cre/ <i>loxP</i> system	117
4.3.2 PCR verification of putative CAS-1116 excision mutants	117
4.3.3 Mobilization of a plasmid borne <i>lpiA-gusA</i> fusion into MUR2122, MUR2128 and MUR2125	120
4.3.4 Expression studies	120
4.3.5 qRT-PCR	122
4.3.5.1 RNA quantification	122
4.3.5.2 Confirmation of primer specificity	123
4.3.5.3 PCR amplification efficiencies for each primer pair	124
4.3.5.4 Calculation of the ratio of gene expression at low pH compared to neutral pH	126
4.4 Discussion	129

CHAPTER 5 *LPiA* PROMOTER LOCALIZATION BY 5' RACE

ANALYSIS.....	135
5.1 Introduction.....	136
5.2 Materials and Methods.....	140
5.2.1 Bacterial strains and plasmids.....	140
5.2.2 RNA extraction	140
5.2.3 5' RACE.....	141
5.2.3.1 First-strand cDNA synthesis	141
5.2.3.2 cDNA purification.....	142
5.2.3.3 Amplification of cDNA	143
5.2.3.4 Poly (A) Tailing of First-strand cDNA	143
5.2.3.5 Nested PCR Amplification of dA-tailed cDNA.....	144
5.2.3.6 Cloning the nested PCR product into pGEM-T easy vector	144
5.2.3.7 DNA sequencing	145
5.2.4 Identification of RpoN enhancer binding protein (EBP) genes in <i>S. medicae</i> WSM419	145
5.3 Results	146
5.3.1 Gene specific primer design for <i>lpiA</i> 5' RACE	146
5.3.2 cDNA synthesis using <i>lpiA</i> 6651 (SP1) primer	146
5.3.3 Nested PCR.....	147
5.3.4 Cloning nested PCR product into pGEM-T easy vector	148
5.3.5 Sequencing analysis to determine the <i>lpiA</i> transcription start site	148
5.3.6 Gene specific primer design for 5' RACE analysis of the <i>acvB</i> gene	148
5.3.7 cDNA synthesis using <i>acvB</i> 4552R (SP1) primer.....	150
5.3.8 Nested PCR.....	152
5.3.9 Cloning and sequencing analysis to determine the <i>acvB</i> transcription start site	154

5.3.10	Identification of the <i>lpiA</i> and <i>acvB</i> promoter	155
5.3.11	Prediction of a possible RpoN enhancer binding protein (EBP) for the expression of <i>lpiA</i> ..	155
5.4	Discussion	168
CHAPTER 6	GENERAL DISCUSSION.....	172
6.1	The main findings of the thesis	174
6.2	A model of <i>lpiA/acvB</i> regulation in <i>S. medicae</i> WSM419	180
6.3	Future research directions	185
BIBLIOGRAPHY	186

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere thanks to my principal supervisor Dr. Wayne Reeve for providing me with the opportunity to complete my PhD study under his professional supervision. His broad knowledge and patience has been a great help to me and enabling me to overcome all the difficulties I have encountered. Secondly, special thanks go to my other supervisors Dr. Ravi Tiwari and Dr. Lambert Bräu. You both have given me excellent advice in my research. I would also like to thank Dr. Graham O'Hara for his kind support to me, as an international student, throughout my PhD studies at Murdoch University.

I would like to thank the members of Center for *Rhizobium* Studies and other scientist friends for sharing their research experiences, providing valuable feedback on my work, and most importantly, always being very encouraging and making me feel motivated to complete my studies.

Also, I would like to thank my non-scientist friends for always providing me great company and allowing me escape from my work to see other beautiful things in life. Finally, I am most of all grateful to my parents for their support and understanding that their daughter has not been able to be with them for so many years. However, my heart has been always there with them!

LIST OF PUBLICATIONS

Conference Contributions

- 1) Rui T, Tiwari, R.P., Bräu, L., O'Hara, G., Howieson, J.G. and Reeve, W.G.(2009): Characterization of the regulatory circuit required for the low pH induction of the adaptive tolerance response protein LpiA in *Sinorhizobium medicae* WSM419. *Proceeding of the 16th International Congress on Nitrogen Fixation*. June 14-19, 2009. Big Sky, Montana.USA
- 2) Rui T, Tiwari, R.P., Bräu, L., O'Hara, G., Howieson, J.G. and Reeve, W.G. (2009): Regulatory circuit of the acid response protein LpiA in *Sinorhizobium medicae* WSM419. *Proceeding of the 15th Australian Nitrogen Fixation Conference*. Nov 8-13, 2009. Margaret River, Western Australia.
- 3) Rui T, Tiwari, R.P., Bräu, L., O'Hara, G., Howieson, J.G. and Reeve, W.G. (2011). Identification of the regulatory genes required for the acid activation of the low pH inducible gene *lpiA* in *Sinorhizobium medicae*. *Proceeding of the 16th International Congress on Nitrogen Fixation*. Nov 27-Dec 1, 2011. 2011. Fremantle, Western Australia.

It is anticipated that two papers will be published from the material presented in this thesis.

ABBREVIATIONS

μ	Micro
$A_{260\text{nm}}$	Absorbance at $\lambda = 260 \text{ nm}$
$A_{280\text{nm}}$	Absorbance at $\lambda = 280 \text{ nm}$
AM3	Antibiotic medium 3
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
bp	Base pair
CATB	Hexadecyltrimethyl ammonium bromide
d	Day(s)
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxynucleoside triphosphates
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
<i>g</i>	Centrifugal force
g	Gram
h	Hour(s)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl β -D-1-thiogalactopyranoside
kb	kilo base pairs
L	Litre
LB	Luria-Bertani
M	Molar
MES	2-(<i>N</i> -morpholino) ethanesulfonic acid
mg	Milligram
min	Minute(s)
MOPS	3-(<i>N</i> -morpholino) propanesulfonic acid

MW	Molecular weight
n	Nano
°C	Degree Celsius
OD _{600nm}	Optical density at $\lambda = 600$ nm
ONPG	ortho-Nitrophenyl- β -galactoside
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
s	Second(s)
SDS	Sodium dodecyl sulfate
TAE	Tris, acetic acid EDTA
Tn5	Transposon 5
Tris	Tris(hydroxymethyl)aminomethane
TY	Tryptone yeast
UV	Ultra-violet
V	Volt(s)
X-Gal	5-bromo-4-chloro-indolyl- β -D-galactopyranoside
X-Glc	5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid

ABSTRACT

The study of low pH responsive gene expression in *Sinorhizobium medicae* has identified many genes that are regulated in response to acidic conditions. One of these, the *lpiA* (*low pH inducible gene A*) gene, is acid-activated at least twenty fold by low pH and not by any other stress. The acid-activated expression of *lpiA* has been shown to require functional FsrR (*fused sensor-regulator*). However, even in the absence of FsrR, there is still a 6-fold acid-induction of *lpiA* revealing that other proteins are required for the acid-activation process. Such regulators could be encoded by *tcsA* (*two component sensor gene A*) and *tcrA* (*two component regulator gene A*) which are located upstream of *fsrR* in the *lpiA* gene region. Furthermore, the identification of a potential sigma factor binding motif upstream of *lpiA* suggested the requirement for RpoN (*RNA polymerase sigma factor for nitrogen metabolism*) for the transcription process. These regulatory proteins may also be required for the transcription of the *acvB* (*acid virulence induced gene B*) since the DNA sequence contains a *lpiA* stop codon that is coupled to the *acvB* start codon. This would infer that these two genes constitute an operon. It was the aim of this work to therefore investigate the additional regulatory aspects critical for the acid activation of *lpiA* and to identify if the *lpiA* and *acvB* genes are co-transcribed in response to acid.

Mutations were constructed by inserting the CAS-1116 cassette into the *tcsA*, *acvB* and *lpiA* genes of *S. medicae* WSM419 to create the mutants MUR2121, MUR2124 and MUR2127, respectively. Each mutant was constructed by single crossover insertional

inactivation to disrupt the targeted gene and to enable expression to be monitored in each background. In addition, mutants were obtained from previous studies carrying single crossover mutations in *tcrA*, *rpoN* and *fsrR*.

Phenotypic characterization of *tcsA*, *tcrA*, *fsrR*, *acvB*, *lpiA* and *rpoN* mutants revealed that these genes are not required for resistance to ZnCl₂ (up to 0.7 mM), CdCl₂ (up to 0.175 mM), CuSO₄ (up to 1.75 mM), NaN₃ (up to 0.175 ml) or ethanol (7%) in comparison to the wild-type WSM419. These genes were also found not to be essential for cell growth in moderately acidic conditions (pH 5.7). However, it was found that a mutation in *lpiA* or *rpoN* slightly decreased the growth rate of the mutants when they were exposed to pH 5.7. Furthermore, in comparison to the wild-type *S. medicae* WSM419, the two mutants were more sensitive to the antibiotic polymyxin B, and were only marginally more sensitive to the antibiotics vancomycin, bacitracin and spectinomycin at pH 5.7. In addition, it was also discovered that the mutation in *rpoN*, but not *lpiA*, caused a symbiotic nitrogen fixation defect for the hosts *Medicago sativa*, *Medicago murex*, *Medicago polymorpha* and *Medicago truncatula*.

All of the above-mentioned mutants described were constructed by the insertion of a CAS-1116 cassette into the targeted gene. The promoterless *gusA* reporter was used to describe enable transcription to be monitored at the point of insertion. β -glucuronidase (GUS) determinations at pH 7.0 and pH 5.7 revealed that all of the CAS-1116 induced fusions, including that to *lpiA*, were constitutively expressed in the conditions tested. This finding indicated the unexpected existence of a strong unidentified promoter in the

cassette sequence upstream of *gusA*. Expression of each gene was therefore measured in cells cultured at pH 7.0 and pH 5.7 using the qRT-PCR technique. This technique confirmed the acid-induction of *lpiA* and revealed that *acvB* was also acid-activated to the same extent. This result reinforced the concept that *acvB* was co-transcribed with *lpiA* in *S. medicae*. In contrast, *tcsA*, *tcrA*, *fsrR* and *rpoN* were shown to be constitutively transcribed with respect to pH.

To investigate the expression of *lpiA* in each mutant, *gusA* negative backgrounds were constructed. The CAS-1116 cassette (flanked by *loxP* sites) was excised from MUR2121 (*tcsA*:CAS-1116), MUR2124 (*acvB*:CAS-1116), MUR2090 (*tcrA*:CAS-1116), MUR2127 (*lpiA*:CAS-1116) and MUR2088 (*rpoN*:CAS-1116) by Cre-mediated recombination. The resulting *S. medicae* mutants MUR2122, MUR2125, MUR2092, MUR2128, and MUR2093 contained Δ CAS-1116 *loxP* mutations in *tcsA*, *acvB*, *tcrA*, *lpiA* and *rpoN*, respectively. To monitor the expression of *lpiA* in these backgrounds, a plasmid borne *lpiA-gusA* fusion located on pWR220-101H was mobilized into the deletion mutant backgrounds MUR2122, MUR2125, MUR2092, MUR2128 and MUR2093 to create the strains MUR2123, MUR2126, MUR2094, MUR2129 and MUR2095, respectively. The expression of the plasmid borne *lpiA-gusA* fusion was examined in each mutant cultured at low and neutral pH. The fusion was partially induced at pH 5.7 in MUR2092 (~6-fold) and MUR2125 (~10-fold) revealing that TcrA and AcvB both act positively to acid activate *lpiA* transcription. The expression of the *lpiA-gusA* fusion in MUR2122 or MUR2092 was totally abolished revealing that both TcsA and RpoN are essential for *lpiA* expression. This result confirmed a role for RpoN and revealed for the first time a

specific sigma factor that is required for pH regulated gene expression in *S. medicae*. Furthermore, the finding provided credence to the suggestion that the putative RpoN binding motif located upstream of *lpiA* may in fact be the *lpiA* promoter.

To further reinforce this concept, the start site of acid-induced transcription was determined using the 5' Rapid Amplification of cDNA ends (RACE) technique. The transcription start site for both *lpiA* and *acvB* was found to occur 206 bp upstream of the *lpiA* start codon. The *rpoN* binding site (5'-TGGCACG-N4-TTGCW-3') was located adjacent to this identified transcription start site but with the bold G and C in the motif positioned -27 and -14 bases upstream of the transcription start site rather than the expected positioning at -24 and -12. RpoN is known to have a requirement for an enhancer binding protein (EBP) to form the open form complex required for transcriptional activation. A bioinformatics search against the WSM419 genome for proteins carrying a sigma54_interaction domain (Pfam accession number: pfam00158) revealed that the gene product of Smed_5956, located upstream of *tcsA*, is a putative EBP. Future work could investigate whether it is this EBP, or another, that is required for the expression of the *lpiA/acvB* operon.

The results presented reveal that the expression of the low pH induced *lpiA/acvB* operon is regulated in a multifaceted way. Interestingly, weak induction of *lpiA* occurs in the absence of TcrA or FsrR. Since induction was totally abolished in a TcsA mutant, this could infer that TcsA is the cognate sensor for each of these two-component regulatory proteins. In addition, full expression of *lpiA* was shown to occur only if the RpoN sigma

factor was present. This is the first report that has detailed a role for the alternative sigma factor RpoN to regulate the transcription of acid-induced genes. The findings of this thesis have been used to construct a model for the *lpiA* acid-induced regulatory circuit and the implications of this model are discussed.